Korean J. Malacol. 38(4): 221-233 2022
https://doi.org/10.9710/kjm.2022.38.4.221

생물정보학을 활용한 멸종위기 연체동물 전사체 서열에 오염된 곰팡이 유전자 발굴

송대권[1,2], 박현진[5], 상민규[3], 박지은[1,2], 정준양[1,2], 홍찬의[1,2], 김용태[1,2], 신현준[1,2], 류자미[1,2], 조용훈[1], 한연수[4], 이용석[1,2,3], 장종수[5]

[1]순천향대학교 자연과학대학 생명과학과, [2]한국자생동물자원활용 융복합연구소, [3]생명자원 바이오빅데이터 분석 및 활용 연구지원센터, [4]전남대학교 농업생명과학대학 식물생명공학부, [5]한국방송통신대학교 농학과

Identification of Fungal Gene Sequence Contamination in Transcriptome Sequence Data of Endangered Molluscs using Bioinformatics

Dae Kwon Song[1,2], Hyeon Jin Park[5], Min Kyu Sang[3], Jie Eun Park[1,2], Jun Yang Jeong[1,2], Chan-Eui Hong[1,2], Yong Tae Kim[1,2], Hyeon Jun Shin[1,2], Ziwei Liu[1,2], Yong Hoon Jo[1], Yeon Soo Han[4], Yong Seok Lee[1,2,3] and Jong Soo Chang[5]

[1]*Department of Biology, College of Natural Sciences, Soonchunhyang University, Asan, Chungnam 31538, Korea*
[2]*Korea Native Animal Resources Utilization Convergence Research Institute*
[3]*research Support Center (Core-Facility) for Bio-Bigdata Analysis and Utilization of Biological Resources*
[4]*Institute of Environmentally-Friendly Agriculture (IEFA), College of Agriculture and Life Sciences, Chonnam National University, Gwangju 61186, Korea*
[5]*Agricultural Science, Korea National Open University Department of Applied Biology,*

ABSTRACT

The amount of data is growing very fast as advances in NGS technology enable the acquisition of large amounts of genome and transcriptome data. Moreover, the accuracy and speed of bioinformatic analysis of NGS data remains of great importance these days. However, the sequence database of mollusks is fall short of other organisms groups, and it thus appears that the annotation results after BLAST analysis are not accurate and reliable due to potential contamination with fungal sequences in mollusks sequence database. In this context, we constructed a BLAST database with 20 species of mollusk unigene sequences and 32 species of fungal sequences derived from previous studies. In order to confirm the contamination of fungal gene sequences in the unigenes of 20 endangered species, bioinformatics analysis was performed using BLAST. It reveals that the NGS sequences of mollusks are mixed with fungal sequences. Taken together, our results suggest that it is essential to reconfirm mollusks sequence information before publication.

Keywords: NGS, Mollusks, Fungal sequences

Received: December 12, 2022; Revised: December 22, 2022; Accepted: December 29, 2022
Corresponding author: Jong Soo Chang
Tel: +82 (02) 3668-4636, e-mail: jschang@knou.ac.kr
1225-3480/24832
Corresponding author: Yong Seok Lee
Tel: +82 (41) 530-3040, e-mail: yslee@sch.ac.kr
This is an Open Access Article distributed under the terms of the Creative Commons Attribution Non-Commercial License with permits unrestricted non-commercial use, distribution, and reproducibility in any medium, provided the original work is properly cited.

서 론

차세대염기서열분석 (NGS; Next Generation Sequencing) 기술이 급속도로 발전함에 따라 저렴한 비용과 짧은 분석시간으로 대량의 유전체 및 전사체 염기서열 데이터의 확보가 가능해졌다 (Metzker, 2010; van Dijk *et al*., 2014). NGS 기술의 발전으로 유전체 프로젝트는 국제컨소시엄에서 국내컨소시엄 단위로 바뀌어 진행되기 시작했으며, 최근에는 미생물, 곰팡이 등 생물의 유전체 길이에 따라 유전자 서열이 짧은 생물은 실험실 단위에

서 수행되고 있다 (Morozova *et al*., 2008; Yang *et al*., 2009; Bang *et al*., 2010; Sang *et al*., 2020). 이에 따라, NCBI의 SRA (Sequence Read Archive) 에 등록되어지는 데이터의 양은 NCBI GenBank의 데이터 양에 비하여 매우 빠른속도로 증가하고 있다 (https://www.ncbi.nlm.nih.gov/genbank/statistics/).

2022년 12월, NCBI에 등록된 연체동물 유전체 정보는 nucleotide 6,135,382건, genome 747건, protein 2,083,175건, gene 646,491건 등의 서열이 등록되어 있으며, 특히 SRA (Sequence Read Archive) 는 41,342 개에 해당하는 Raw data 형태의 NGS 데이터가 등록되어 있다. 그 중 가장 많은 서열을 가진 *Crassostrea gigas* 가 4,974건으로 다양한 연구가 수행되고 있었으며 (Sang *et al*., 2020), *Crassostrea virginica* 가 2,350건으로 두 번째로 많았다. 이러한 NGS 데이터는 크게 WGS (Whole Genome Shotgun) 과 RNAseq (transcriptome) 서열로 나눌 수 있으며, NGS 데이터에 대한 생물정보학적 분석의 정확성과 속도는 최근에도 매우 중요하게 다루어지고 있다.

균류는 에너지의 순환과정에서 핵심적인 기능과 역할을 수행함으로써 생태계를 안정적으로 유지하게 하는 필수적인 생물 종으로 종 다양성이 매우 뛰어난 것으로 알려져 있다 (Blackwell, 2011). 육지에서 주로 서식하는 것으로 알려져 있던 균류가 해양환경에서도 발견되어, 균류의 다양한 서식지에 대한 연구가 다수 진행되고 있으며 (Hyde *et al*., 1998; Kim and Harvell, 2004; Godinho *et al*., 2019), 바다의 기역별로 해양 균류에 대한 분류학 연구도 시도되었다 (Kohlmeyer and Volkmann-Kohlmeyer, 1991).

1954년 Davis *et al*., 에 의해 조개와 굴의 유생에 감염되어진 곰팡이에 대하여 언급한 것에 착안하여 (Davis *et al*., 1954), Zvereva and Vysotskaya (2005) 에서 언급되어 동해 (East sea) 의 이매패류에 감염되어진 32종의 곰팡이의 (Zvereva and Vysotskaya, 2005) 서열을 모두 모아 데이터베이스를 만들고 본 연구진에 의해 발굴된 20종의 연체동물 unigene 서열을 query로 하여 발굴된 unigene 중 어느 정도의 서열이 fungi의 서열로 판단되어질 수 있을지 알아보았다. 본 연구는 각종 연체동물의 NGS 서열에 fungi 가 혼합되어 있을 수 있는 가능성을 알아보고 이를 기반으로 하여 추후 진행되어질 많은 유전체 및 전사체의 생물정보학적 분석의 정확도를 높일 수 있는 기반을 마련하고자 수행되었다.

재료 및 방법

1. 연체동물 감염가능성을 가진 곰팡이의 유전자원 확보 및 DB구축

NCBI Pubmed와 Google Scholar 검색을 통하여 molluscs (or mollusks), fungi, infection을 키워드로 하여 참고문헌들을 검색하여 연체동물에 감염될 수 있는 곰팡이 32종을 확인하였다. 확인되어진 32종의 곰팡이의 유전자원을 확보하기 위해 NCBI database 에서 *Acremonium*, *Alternaria*, *Aspregillus*, *Aureobasidium*, *Chaetomium*, *Cladosporium*, *Penicillium*, *Myxotrichum*, *Xanthiomyces* 속에 속하는 Filamentous fungi의 nucleotide 서열과 amino acids 서열을 확보하고, BLAST+ software의 makeblastdb 프로그램을 사용하여 BLAST가 가능한 database로 구축하였다.

2. 멸종위기 연체동물 20종의 unigene 서열 확보

본 연구진에 의해 구축된 redlist 데이터베이스 (http://redlist.sch.ac.kr) 에 등록이 되어있는 멸종위기 연체동물 20종의 unigene 서열들을 모두 다운로드 하여 확보하였다. 사용되어진 unigene 서열들은 모두 NGS로 생산된 Raw reads를 Cutadapt program (Martin, 2011) 을 사용하여 Adapter sequence들을 제거한 뒤, sickle program을 사용하여 low-quality region (quality score 20 이하) 을 제거한 clean reads (high-quality reads) 들을 확보하였다. 이후 Trinity program (Grabherr *et al*., 2011) 을 이용하여 *De novo* assembly (k-mer size 25, 최소 contig 길이 200 bp

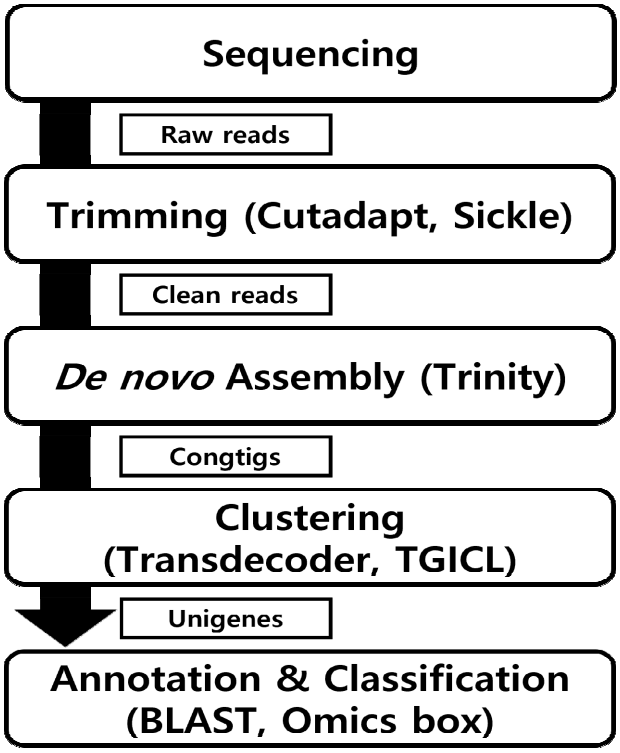

Fig. 1. Bioinformatic analysis pipeline for unigene acquisition of 20 mollusk species.

Table 1. Status of established genetic resource database of filamentous fungi

No.	Species of fungi	Nucleotide Sequence 수	Protein Sequence 수
1	*Acremonium* sp.	1,231	164
2	*Alternaria alternata* (*Fresen*) *Keissl*	57,910	44
3	*Aspregillus candidus* Link	10,521	19,379
4	*Aspregillus glaucus*	11,617	22,579
5	*Aspregillus janus* Raper et Thom	36	16
6	*Aspregillus nidulans* (Eidam) G. Winter	33,204	72,434
7	*Aspregillus niger* Tiegh	92,438	558,049
8	*Aspregillus ochraceus* K. Wilh	710	176
9	*Aspregillus parasiticus* Speare	27,613	25,471
10	*Aspregillus speluneus* Raper et Fennell	33	10
11	*Aspregillus sulphureus* Thom et Church	30	12
12	*Aspregillus versicolor* (Vuill.) Tirab	14,682	26,665
13	*Aspergillus* sp.	8,515	30,359
14	*Aureobasidium pullulans* (de Bary) G. Arnaud	129,250	675,075
15	*Cladosporium cladosporioides* (Fresen.) B. de Vries	3,806	1,013
16	*Cladosporium sphaerospermum* Penzes	1,512	584
17	*Cladosporium oxysporum* Berk. et M.A. Curtis	288	47
18	*Penicillium aurantiogriseum* Dierckx	77,436	126
19	*Penicillium citrinum* Thom	2,815	5,173
20	*Penicillium citreonigrum* Dierckx	236	75
21	*Penicillium commune* Thom	676	123
22	*Penicillium decumbens* Thom	326	17,245
23	*Penicillium griseofulvum* Dierckx	10,261	19,376
24	*Penicillium jensenii* W. Zalessky	150	7
25	*Penicillium* sp.	14,972	58,375
26	*Dichotomopilus funicola*	174	16
27	*Chaetomium globosum* Kunze	16,269	35,118
28	*Chaetomium megalocarpum* Bainier	72	19
29	*Chaetomium spiculipilium* L.M. Ames	10	4
30	*Xanthiomyces spinosus*	12	4
31	*Chaetomium* sp.	1,717	21,808
32	*Myxotrichum chertarum* Kunze	6	0

이상) 과정을 수행하였다. assembly 결과로 얻은 contigs를 TGICL (TIGR Gene Indices Clustering tools) software (Pertea *et al*., 2003) 를 사용하여 clustering (identity 94%, overlap 30 bp)을 진행하여 확보되어진 서열들로 대부분 publish 된 바 있다 (Fig. 1).

3. BLAST를 활용 한 멸종위기 연체동물 20여종의 unigene 과 확보된 곰팡이 서열의 비교 및 R-package 를 활용한 데이터 분석

선행된 연구에서 확보한 20종의 연체동물의 unigene data (Table. 2) 를 NCBI BLAST+ software (Altschul *et al*., 1990; Camacho *et al*., 2009) 로 데이터베이스화 하여, 연체

Table 2. List of species of mollusks used in the study include unigene information

No.	Species	Class	국명	Unigene 수	Publish
1	*Aegista* (Aegista) *chejuensis*	Gastropoda	제주배꼽달팽이	34,357	(Kang *et al.*, 2016) (Park *et al.*, 2018)
2	*Aegista* (Plectotropis) *quelpartensis*	Gastropoda	제주배꼽털달팽이	257,005	(Kang *et al.*, 2016) (Jeong *et al.*, 2020)
3	*Karaftohelix adamsi*	Gastropoda	울릉도달팽이	13,753	Unpublished
4	*Koreanohadra kurodana*	Gastropoda	북한산달팽이	27,835	(Kang *et al.*, 2016)
5	*Incilaria fruhstorferi*	Gastropoda	산민달팽이	50,230	(Patnaik *et al.*, 2019) (Sang *et al.*, 2017)
6	*Satsuma myomphala*	Gastropoda	거제외줄달팽이	12,191	(Kang *et al.*, 2017) (Sang *et al.*, 2016) (Sang *et al.*, 2020)
7	*Charonia sauliae*	Gastropoda	나팔고둥	35,951	(Chung *et al.*, 2020)
8	*Clithon retropictus*	Gastropoda	기수갈고둥	31,608	(Park *et al.*, 2016) (Jeong *et al.*, 2021) (Hong *et al.*, 2021)
9	*Ellobium chinense*	Gastropoda	대추귀고둥	32,510	(Kang *et al.*, 2018) (Park *et al.*, 2020)
10	*Koreanomelania nodifira*	Gastropoda	염주알다슬기	49,033	Unpublished
11	*Mirus junensis*	Gastropoda	두타산입술대고둥아재비	53,775	(Min *et al.*, 2018)
12	*Pugilina ternatanus*	Gastropoda	털탑고둥	222,989	Unpublished
13	*Thais bronni*	Gastropoda	두드럭고둥	40,331	Unpublished
14	*Anodonta woodiana*	Biavalvia	펄조개	351,660	Unpublished
15	*Arca boucardi*	Biavalvia	긴네모돌조개	244,863	Unpublished
16	*Crassostrea echinata*	Biavalvia	가시굴	232,490	Unpublished
17	*Cristaria plicata*	Biavalvia	귀이빨대칭이	374,794	(Patnaik *et al.*, 2016) (Chung *et al.*, 2017)
18	*Lamprotula coreana*	Biavalvia	두드럭조개	71,798	Unpublished
19	*Lamprotula leai*	Biavalvia	곳체두드럭조개	50,607	Unpublished
20	*Pisidium coreanum*	Biavalvia	산골조개	451,772	(Jeong *et al.*, 2015)

동물 unigene 데이터베이스를 제작하였다. 이후, 연체동물에 감염되는 것으로 알려진 32종의 fungi nucleotide sequence (Table. 2) (Zvereva *et al.*, 2005) 를 연체동물 unigene 데이터베이스에 BLASTn (E-value ≤ $1e^{-5}$) 을 수행하였다. 또한, 20종의 연체동물 unigene data를 NCBI BLAST+ software로 데이터베이스화한 32종의 fungi protein sequence (Table. 3) 에 BLASTx (E-value ≤ $1e^{-5}$) 를 수행하였다. BLAST를 통한 서열 비교 분석은 순천향대학교 생명자원 바이오빅데이터 분석 및 활용 연구지원센터의 특화 데이터베이스를 활용한 서열 비교 분석 서비스 (Z-202212303989)를 활용하여 분석하였다. BLAST를 통해 생성된 결과는 R program package인 dplyr (Wickham *et al.*, 2020) 를 사용하여 fungi sequence의 summary를 정형화한 데이터프레임에 R program의 left_join function을 사용하여 merge하

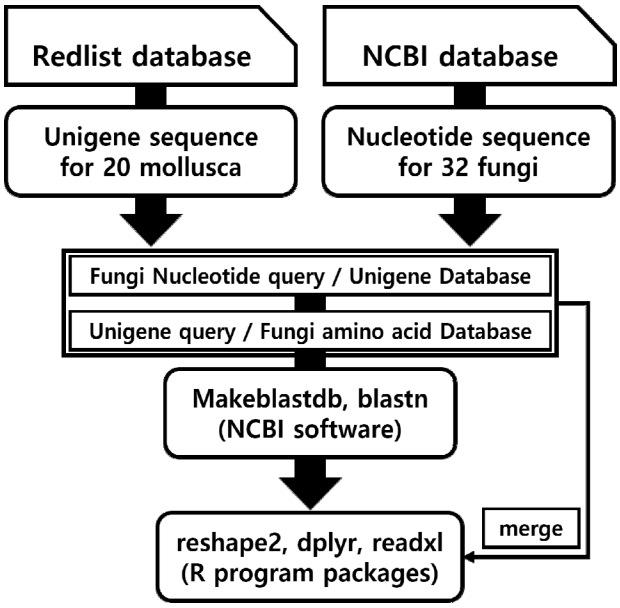

Fig. 2. Bioinformatic analysis process of mollusk unigene and fungal genetic resources using BLAST and R package.

였다. 제작된 데이터프레임은 R program package인 reshape2 의 dcast function (Wickham, 2007) 을 사용하여 pivot table을 작성하였다 (Fig. 2).

결과 및 고찰

1. 연체동물 감염가능성을 가진 곰팡이의 유전자원 확보 결과

NCBI Pubmed와 Google Scholar 검색을 통하여 "molluscs" ("mollusks"), "fungi" 의 키워드 검색결과 약 15,000편의 논문이 검색되었으나 대부분 연체동물의 전사체에서 항균 효과를 가진 유용유전자를 찾기 위한 문헌들이 주로 검색되었으며, 본 연구진이 확보하고자 하는 연체동물 감염관련하여 곰팡이의 종명이 명기되어 있는 연구는 거의 수행되어지지 않았다. 그 중 일부 논문은 단편적으로 몇몇 생물에서 분리 또는 발견된 곰팡이에 대한 논문들 이었다. 그 중 Zvereva *et al*., (2005) 등에 의해 연구된 논문에서 연체동물에 감염되어진 35종의 곰팡이 종명이 나열되어 있었다 (Zvereva *et al*., 2005). 그 중 3종은 서열분석이 전혀 이루어 지지 않았다. 이에 본 연구진은 32종의 곰팡이에 대한 유전자원을 확보하였으며, 516,805개의 nucleotide sequence와 1,567,738개의 protein sequence 모두 makeblastdb 를 활용하여 BLAST가 가능한 데이터베이스로 구축하였다. 구축되어진 BLAST 가능 DB에 포함된 32종의 곰팡이 중 nucleotide가 가장 많이 등록되어진 종은 *Aureobasidium pullulans* 로 129,250개가 등록되어 있었으며, 아미노산 서열이 가장 많이 등록되어진 종은 *Aspregillus niger* 로 558,049 개의 서열이 등록되어 있었다 (Table 1).

2. 확보되어진 멸종위기 연체동물 20여종의 unigene 서열

Redlist 서버 (http://redlist.sch.ac.kr) 에서 다운로드하여 확보된 총 20종의 unigene 데이터는 복족강 13종 과 부족강 7종으로 구성되어 있었으며, unpublished 데이터는 9종이 포함되어 있었다. 본 연구에 사용된 연체동물의 총 unigene 데이터는 2,639,552개 였다. 복족강에 속하는 연체동물 13종의 경우 861,568개의 unigene 서열이 확보되었으며, 이 중 제주배꼽털달팽이 (*Aegista* (*Plectotropis*) *quelpartensis*) 가 257,005개로 가장 많았다. 부족강에 속하는 연체동물 7종의 경우 1,777,984개의 unigene 서열이 확보되었는데, 산골조개 (*Pisidium coreanum*) 가 451,772개로 가장 많았다 (Table 2).

3. 멸종위기 연체동물 20여종의 unigene 과 확보된 곰팡이 서열의 비교 결과

20종의 연체동물 unigene 서열 중 곰팡이의 서열이 존재하는지 알아보기 위하여 구축되어진 32종의 fungi nucleotide sequence 를 query로 구축한 연체동물 unigene 데이터베이스에 BLASTn 한 결과, 긴네모돌조개 (*Arca boucardi*) 의 unigene에 29,903개로 가장 많은 fungi가 매칭되었으며, 털탑고둥 (*Pugilina ternatanus*) 5,403개, 펄조개 (*Anodonta woodiana*) 3,971개, 기수갈고둥 (*Clithon retropictus*) 2,393개 순으로 매칭되었다. 가장 적은 숫자의 unigene 이 fungi 서열과 매칭 된 종은 거제외줄달팽이 (*Satsuma myomphala*) 로 39개가 매칭되었다. 각 연체동물의 unigene 서열과 BLASTn으로 alignment되어진 fungi를 백분율로 환산한 결과, 긴네모돌조개 (*Arca boucardi*) 가 12.21%로 가장 많은 fungi서열이 unigene 서열에 매칭되었으며, 기수갈고둥 (*Clithon retropictus*) 이 7.57%, 두드럭고둥 (*Thais bronni*)이 4.9%, 털탑고둥 (*Pugilina ternatanus*) 이 2.42% 의 순으로 나타났다 (Table. 3, Fig. 3).

BLASTn 결과를 종 수준에서 확인했을 때, 총량을 기준으로 곰팡이는 *Alternaria alternata*가 20.07%로 가장 많이 매칭되었으며, *Penicillium* sp. (16.77%), *Aureobasidium pullulans* (14.13%), *Aspergillus niger* (11.64%) 순으로 나타났다. *Alternaria alternata*가 가장 많이 매칭된 연체동물은 긴네모돌조개 (*Arca boucardi*) 와 기수갈고등 (*Clithon retropictus*), 털탑고둥 (*Pugilina ternatanus*) 이 가장 많이 매치되었으며, *Penicillium* sp.가 많이 매칭된 연체동물은 펄조개 (*Anodonta woodiana*) 로 확인되었다. *Aureobasidium*

Table 3. Alignment results of filamentous fungus sequences against mollusk unigenes database using BLASTn

Species of mollusca	Total amount Unigene	Alignment fungi	Percentage (%)
Cristaria plicata	374,794	336	0.09
Pisidium coreanum	451,772	688	0.15
Lamprotula coreana	71,798	121	0.17
Lamprotula leai	50,607	117	0.23
Koreanohadra kurodana	27,835	75	0.27
Aegista (Plectotropis) *quelpartensis*	257,005	786	0.31
Satsuma myomphala	12,191	39	0.32
Ellobium chinense	32,510	118	0.36
Incilaria fruhstorferi	50,230	224	0.45
Mirus junensis	53,775	346	0.64
Crassostrea echinata	232,490	1,764	0.76
Aegista (Aegista) *chejuensis*	34,357	355	1.03
Anodonta woodiana	351,660	3,971	1.13
Karaftohelix adamsi	13,753	164	1.19
Koreanomelania nodifira	49,033	781	1.59
Charonia sauliae	35,951	693	1.93
Pugilina ternatanus	222,989	5,403	2.42
Thais bronni	40,331	1,976	4.90
Clithon retropictus	31,608	2,393	7.57
Arca boucardi	244,863	29,903	12.21

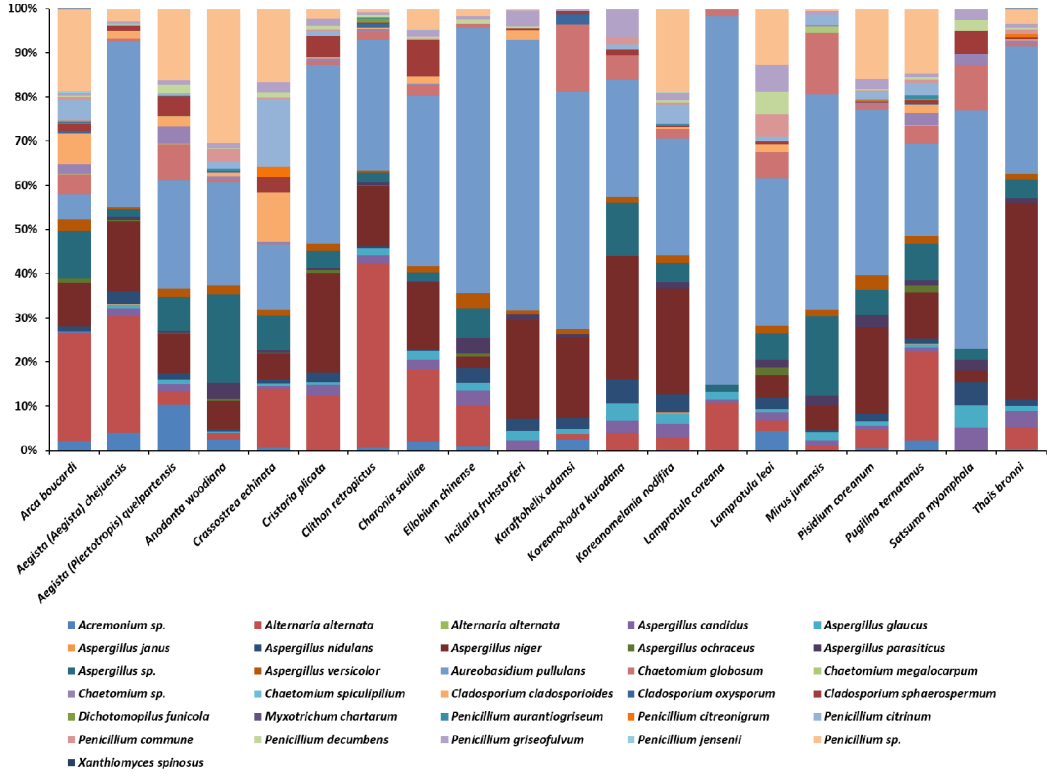

Fig. 3. Types of filamentous fungi (Species level) identified in mollusk unigene using BLASTn.

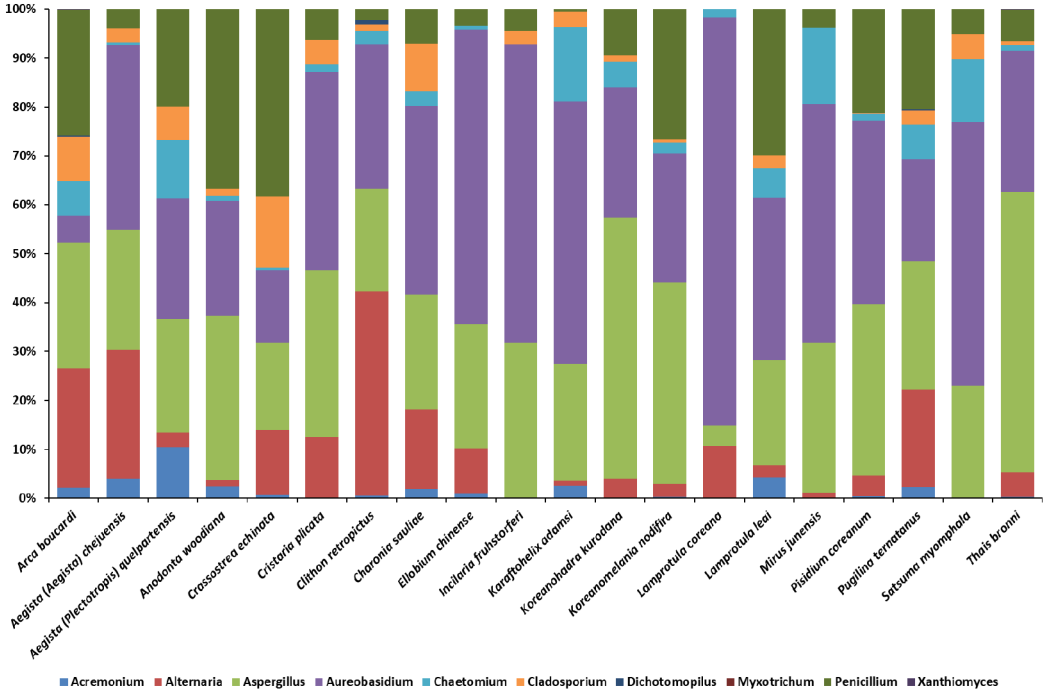

Fig. 4. Types of filamentous fungi (Genus level) identified in mollusk unigene using BLASTn.

pullulans 는 총량을 기준으로 세 번째로 많이 매칭된 곰팡이지만 가장 많은 연체동물 종에 매치가 되었다. 연체동물에 가장 많이 매칭된 곰팡이 *Alternaria alternata*는 다양한 해양 식물과 갑오징어와 같은 무척추 동물의 표면과 내부조직에서 추출하여 해양 동물 유래 균류의 2차 대사산물에 대한 연구가 시도되었음을 확인하였다 (Zhang *et al*., 2009; Ramasamy *et al*., 2011).

BLASTn 결과를 속 수준에서 확인했을 때, 총량을 기준으로 곰팡이는 Aspergillus가 27.62% 이상 많이 매칭되었으며 Penicillum (23.51%), Alternaria (20.07%), Aureobasidium (14.13%) 순으로 나타났다. Aspergillus는 긴네모돌조개 (*Arca boucardi*), 북한산달팽이 (*Koreanohadra kurodana*), 염주알다슬기 (*Koreanomelania nodifira*), 털탑고둥 (*Pugilina ternatanus*) 에서 가장 많이 매칭되었다. Penicillum 은 펄조개 (*Anodonta woodiana*), 가시굴 (*Crassostrea echinate*) 에서 가장 많이 매칭되었다. 가장 많은 연체동물에 매칭된 Aureobasidium은 12종의 연체동물에 매칭된 것을 확인할 수 있었다. BLASTn결과에서 두 번째로 많이 매칭되어진 Penicillium 속은 수생태계에서 무척추동물과 수초 등에 서식하고 있을 가능성이 높은 곰팡이이며, 식품으로의 가치가 있는 해양 연체동물과 Penicillium 곰팡이에 대한 미생물학적 분석 연구가 진행되었다 (Bagy *et al*., 1992; Udo *et al*., 2022).

NCBI의 BLAST+ software 프로그램을 사용하여 32종의 fungi nucleotide sequence를 20종의 연체동물 unigene database에 BLASTx을 수행한 결과, 산골조개 (*Pisidium coreanum*) 의 unigene에 35,900개로 가장 많은 fungi가 매칭되었으며, 귀이빨대칭이 (*Cristaria plicata*) 21,660개, 가시굴 (*Crassostrea echinate*) 에 19,898개, 펄조개 (*Anodonta woodiana*) 19,600개의 순으로 매칭되었다. 울릉도달팽이 (*Karaftohelix adamsi*) 가 2,192개로 가장 적은 fungi가 매칭되었다. 각 연체동물의 unigene 서열에 BLASTx로 alignment 되어진 fungi를 백분율로 환산한 결과, 산골조개 (*Pisidium coreanum*) 가 15.01%로 가장 많은 fungi 서열이 unigene 서열에 매칭되었으며, 귀이빨대칭이 (*Cristaria plicata*) 9.06%, 가시굴 (*Crassostrea echinate*)이 8.32%, 펄조개 (*Anodonta woodiana*) 8.2% 순으로 나타났다 (Table. 4, Fig. 5).

BLASTx 결과를 종 수준에서 확인했을 때, 총량을 기준으로 곰팡이는 *Aureobasidium pullulans*가 16.94%로 가장 많이 매칭되었으며, *Alternaria alternata* (14.53%), *Aspergillus niger* (12.42%), *Chaetomium* sp. (10.07%) 순으로 나타났다. BLASTx에서는 각 연체동물마다 매칭되어진 fungi의 우점량이 비슷하였으며, 20종의 연체동물에서 *Aureobasidium pullulans*이 가장 많이 매칭된 것을 확인하였다 (Fig. 5). *Aureobasidium pullulans* 의 경우 바이오촉매 중 하나인 셀룰로오스 분해 효소의 생산에 대한 연구가 시도되었으며 천연

Table 4. Alignment results of filamentous fungus sequences against mollusk unigenes database using BLASTx

Species of mollusca	Total amount Unigene	Alignment fungi	percentage (%)
Karaftohelix adamsi	13,753	2,192	0.92
Satsuma myomphala	12,191	2,441	1.02
Lamprotula leai	50,607	5,343	2.23
Koreanohadra kurodana	27,835	6,435	2.69
Clithon retropictus	31,608	6,983	2.92
Aegista (Aegista) *chejuensis*	34,357	7,081	2.96
Charonia sauliae	35,951	7,129	2.98
Ellobium chinense	32,510	7,160	2.99
Lamprotula coreana	71,798	8,013	3.35
Incilaria fruhstorferi	50,230	9,914	4.15
Thais bronni	40,331	10,199	4.27
Koreanomelania nodifira	49,033	11,086	4.64
Mirus junensis	53,775	11,376	4.76
Pugilina ternatanus	222,989	11,948	5.00
Aegista (Plectotropis) *quelpartensis*	257,005	15,743	6.58
Arca boucardi	244,863	18,996	7.94
Anodonta woodiana	351,660	19,600	8.20
Crassostrea echinata	232,490	19,898	8.32
Cristaria plicata	374,794	21,660	9.06
Pisidium coreanum	451,772	35,900	15.01

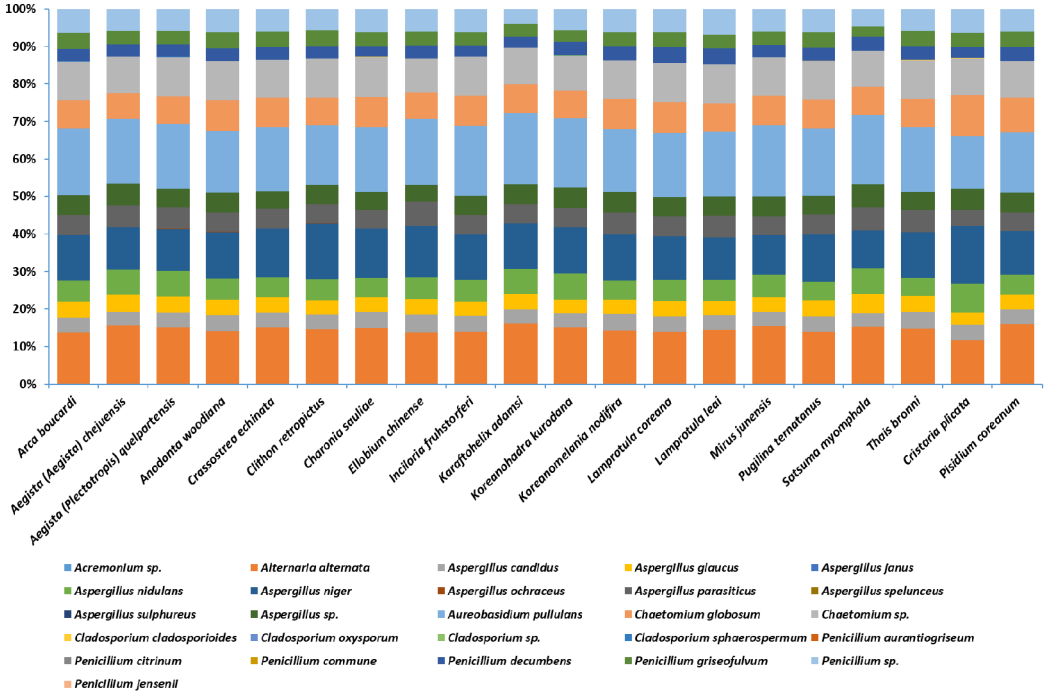

Fig. 5. Types of filamentous fungi (Species level) identified in mollusk unigene using BLASTx.

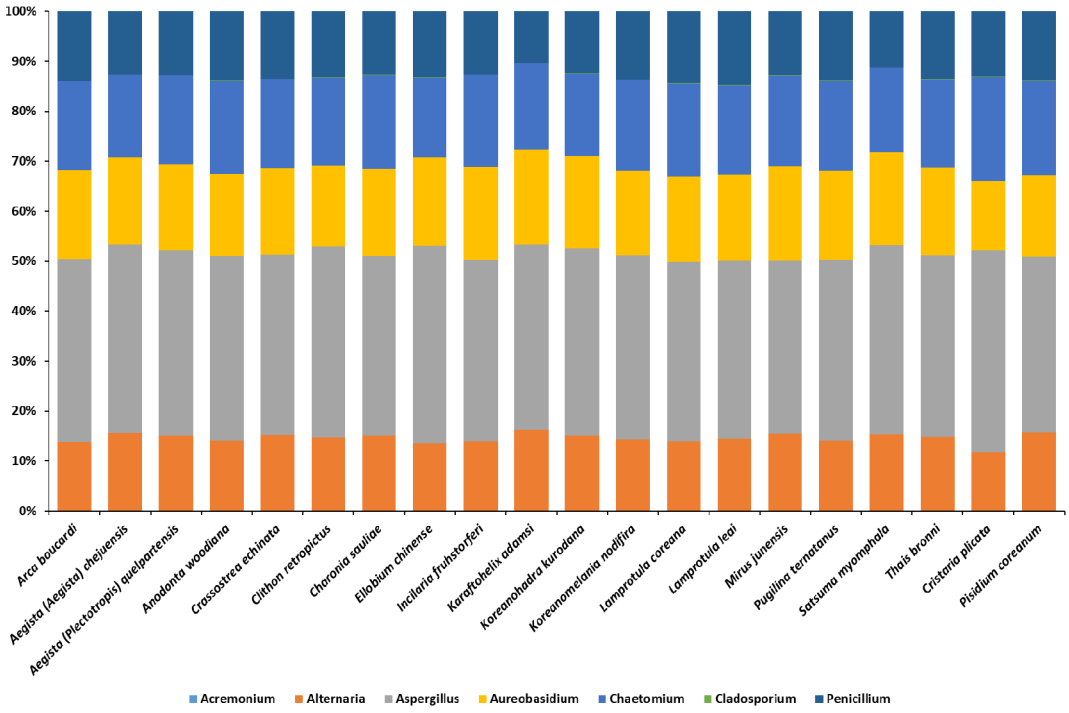

Fig. 6. Types of filamentous fungi (Genus level) identified in mollusk unigene using BLASTx.

물 산업과 관련한 연구가 시도되었다 (Faulkner, 2002; Barzkar and Sohail, 2020).

BLASTx 결과를 속 수준에서 확인했을 때, 곰팡이는 Aspergillus 속이 36.74%로 가장 많이 매칭되었으며 Chaetomium (18.31%), Aureobasidium (16.94%), Alternaria (14.53%), Penicillium (13.41%), Cladosporium (0.06%) 속 순으로 나타났다. Acremonium은 0.02%로 매우 적은 서열이 매칭된 것을 확인하였다 (Fig. 6). 두 번째로 많이 매칭 되어진 Chaetomium 속은 토양 미생물 군집으로 특별한 기능의 구조와 구성을 조절하는 것으로 연구되어 있으며, 국내에서는 제주도 근해의 해수와 해저 퇴적물에서 추적 배양한 연구 등이 수행되었다 (Blunt *et al*., 2012; Osman, 2013; Ryu, 2017). BLASTn의 결과와 BLASTx의 결과 모두 가장 많이 매칭되어진 Aspergilus 속은 항균 활성과 효소 억제 활성에 관여하는 해양 유래 곰팡이로 이매패류 등에 공생하는 것으로 연구되어 있다 (Golubic *et al*., 2005; An *et al*., 2021; Azeem *et al*., 2022).

Kang *et al*., (2019) 에 의하면 연체동물 전사체 분석으로 생성된 1,000개의 유전자 서열로 BLASTx를 수행한 unigene의 annotation 결과에서, BLAST 결과가 없는 서열이 상당히 존재하고 있었으며 annotation이 되어지지 않은 결과를 해석함에 있어서 아직 연체동물의 서열 데이터베이스가 부족한 점이 확인되었다 (Kang *et al*., 2019). 또한 BLAST DataBase (NCBI nr, PANM 등) 를 활용하여 분석할 시, 공생, 기생 및 감염체 관련 서열에 대해 분석되지 않는다. 이는 SRA에 등록되어져 있는 감염체 서열들이 분석에 활용되지 못하기 때문인 것으로 사료된다. 그러므로 연체동물의 감염체에 대한 연구가 많이 수행되어 관련 종의 목록이 확보되어질 필요성이 있으며, 감염체에 대한 유전체, 전사체 연구도 많이 수행되어져야 명확한 정보의 분석이 가능해 질 것으로 판단된다.

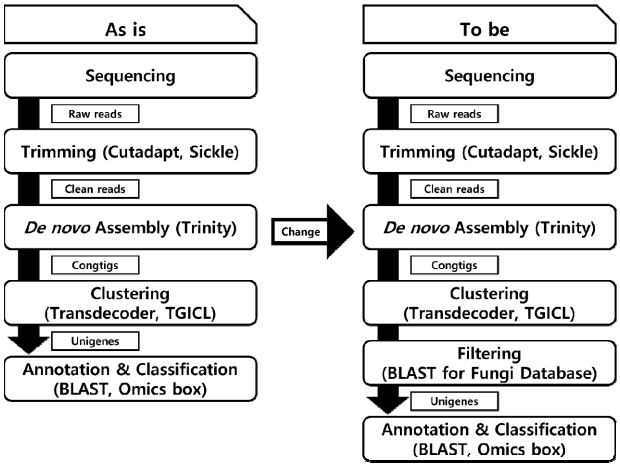

Fig. 7. Improved transcriptome analysis bioinformatics pipeline to remove infected fungal sequences.

그리고 연체동물의 유전체 및 전사체 데이터에서 반드시 곰팡이 서열을 필터링 알고리즘을 추가하여 unigene의 퀄리티를 끌어올릴 필요성이 요구된다고 판단된다. 따라서 연체동물 서열에서 곰팡이 서열을 필터링 할 수 있는 데이터베이스의 제작과 동시에 주기적으로 데이터베이스의 업데이트가 필요하며, 이전에 수행되었던 유전체 및 전사체 데이터 분석의 파이프라인을 업데이트하여 새로운 분석 파이프라인이 필요하다고 사료된다 (Fig. 7.). 또한, NGS의 발달에 의해 많은 데이터가 확보되어지고 있지만 연체동물 내 공생, 기생 및 감염체에 대한 정보는 현저히 부족하므로 관련 기초연구가 많이 수행되어질 필요가 있다고 판단된다.

요 약

1. 본 연구는 생물정보학을 활용하여 멸종위기 연체동물의 전사체 분석결과인 unigene 서열에 곰팡이 유래 유전자서열이 포함되어 있을 가능성을 구명하고자 수행되었다.

2. 분석에 활용된 멸종위기 연체동물 unigene 서열들은 redlist db (http://redlist.sch.ac.kr)에서 다운로드 하여 사용하였다.

3. 분석에 활용되어진 32종의 곰팡이 서열은 Zvereva *et al*., (2005) 등에 의해 연구된 논문에서 연체동물에 감염되어진 35종의 곰팡이중 NCBI 에 유전정보가 등록되어진 32종을 대상으로 하였다.

4. 멸종위기 연체동물의 전사체와 32종의 곰팡이 유전자의 BLAST 분석결과 unigene 으로 분석되어진 멸종위기 연체동물의 전사체에 상당량이 오염되어 있음을 확인하였다.

5. 확보된 곰팡이 유전자원을 토대로 연체동물 유래 곰팡이 데이터베이스의 제작을 통해 clustering되어진 연체동물 unigene에서 연체동물 유래 곰팡이 서열을 필터링 할 수 있는 분석 파이프라인이 필요하다고 판단되어 기존의 파이프라인 개선안을 도출하였다 (Fig. 7).

6. 연체동물 뿐만 아니라 다양한 동물들에 공생 및 기생하는 곰팡이 서열 데이터베이스의 작성이 필요하며, 이를 기반으로 이전보다 더욱 퀄리티가 높은 unigene을 생성하는데 기여하여 유전체 및 전사체 서열에 대한 annotation 분석의 정확도를 향상시키고 시간을 단축시킴으로써 연체동물을 연구하는 연구자들에게 유용하게 이용될 것으로 사료된다.

사 사

이 논문은 2021년도 한국방송통신대학교 학술연구비 지원을 받아 작성된 것임.

REFERENCES

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *Journal of Molecular Biology*., **215**(3): 403-410.

An, C.-L., Kong, F.-D., Li, Y., Ma, Q.-Y., Xie, Q.-Y., Yuan, J.-Z., Zhou, L.-M., Dai, H.-F., Yu, Z.-F., and Zhao, Y.-X. (2021) Asperpenes D and E from the fungus *Aspergillus* sp. SCS-KFD66 isolated from a bivalve mollusk, Sanguinolaria chinensis. *Journal of Asian Natural Products Research,* **23**(2): 117-122.

Azeem, H.H.A.-E., Osman, G.Y., El-Seedi, H.R., Fallatah, A.M., Khalifa, S.A.M., and Gharib, M.M. (2022) Antifungal Activity of Soft Tissue Extract from the Garden Snail *Helix aspersa* (Gastropoda, Mollusca). *Molecules,* **27**(10): 3170.

Bagy, M.M.K., Khallil, A.M., and Obuid-Allah, A.H. (1992) Fungi inhabiting some aquatic macro-invertebrates and water plants of the Nile at Egypt. *Zentralblatt für Mikrobiologie*, **147**(7): 459-475.

Bang, I.S., Han, Y.S., Lee, J.S., and Lee, Y.S. (2010) Current Status of Genome Research in Phylum Mollusks. *The Korean Journal of Malacology,* **26**(4): 317-326.

Barzkar, N., and Sohail, M. (2020) An overview on marine cellulolytic enzymes and their potential applications. *Applied Microbiology and Biotechnology,* **104**(16): 6873-6892.

Binder, M., Hibbett, D.S., Wang, Z., and Farnham, W.F. (2006) Evolutionary relationships of Mycaureola dilseae (Agaricales), a basidiomycete pathogen of a subtidal rhodophyte. *American Journal of Botany*., **93**(4): 547-556.

Blackwell, M. (2011) The Fungi: 1, 2, 3 … 5.1 million species? *American Journal of Botany,* **98**: 426-438.

Blunt, J.W., Copp, B.R., Keyzers, R.A., Munro, M.H., and Prinsep, M.R. (2012) Marine natural products. *Natural Product Report*., **29**(2): 144-222.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T.L. (2009) BLAST+: architecture and applications. *BMC Bioinformatics,* **10**: 421.

Chung, J.M., Hwang, H.J., Min, H.R., Park, J.E., Sang, M.K., Park, S.Y., Park, Y.S., Noh, M.Y., Jo, Y.H., Han, Y.S., Lee, J.S., Park, S.H., Kang, S.W., Kang, C.S., and Lee, Y.S. (2017) Molecular Phylogenetic Analysis based on Metallothionein Gene Sequence of an Endangered Species *Cristaria plicata* in Korea. *The Korean Journal of Malacology,* **33**(1): 35-40.

Chung, J.M., Hwang, H.J., Sang, M.K., Park, J.E., Song, D.K., Jeong, J.Y., Jo, Y.H., Han, Y.S., Park, H.S., Lee, J.S., Shin, E.H., Park, S.Y., and Lee, Y.S. (2020) Phylogenetic analysis of *Charonia sauliae* inferred from Aquaporin gene. *The Korean Journal of Malacology,* **36**(1): 35-43.

Davis, H.C., Loosanoff, V.L., Weston, W.H., and Martin, C. (1954) A Fungus Disease in Clam and Oyster Larvae. *Science,* **120**(3105): 36-38.

Faulkner, D.J. (2002) Marine natural products. *Natural Product Report*., **19**(1): 1-48.

Godinho, V.M., de Paula, M.T.R., Silva, D.A.S., Paresque, K., Martins, A.P., Colepicolo, P., Rosa, C.A., and Rosa, L.H. (2019) Diversity and distribution of hidden cultivable fungi associated with marine animals of Antarctica. *Fungal Biology,* **123**(7): 507-516.

Golubic, S., Radtke, G., and Le Campion-Alsumard, T. (2005) Endolithic fungi in marine ecosystems. *Trends Microbiol.,* **13**(5): 229-235.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., and Regev, A. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*., **29**(7): 644-652.

Hong, C.E., Sang, M.K., Song, D.K., Park, J.E., Jeong, J.Y., Kim, Y.T., Shin, H.J., Liu, Z., Hwang, H.J., Park, S.Y., Kang, S.W., Lee, J.S., Jung, K.Y., and Lee, Y.S. (2021) Molecular Phylogenetic Studies of *Clithon retropictus* Using Arginine Kinase Gene Sequence. *The Korean Journal of Malacology,* **37**(3): 125-132.

Hyde, K.D., Jones, E.B.G., Leaño, E., Pointing, S.B., Poonyth, A.D., and Vrijmoed, L.L.P. (1998) Role of fungi in marine ecosystems. *Biodiversity & Conservation,* **7**: 1147-1161.

Jeong, J.E., Kang, S.W., Hwang, H.J., Park, S.Y., Patnaik, B.B., Kim, C.M., Kim, S.O., Nam, M.M., Lee, J.B., Wang, T.H., Park, E.B., Yi, S.S., Han, Y.S., Lee, J.S., Park, H.S., and Lee, Y.S. (2015) Expressed sequence tag analysis and annotation of genetic information from the freshwater clam, *Pisidium* (Neopisidium) *coreanum* endemic to Korea. *Genes & Genomics,* **37**: 1041-1049.

Jeong, J.Y., Hwang, H.J., Chung, J.M., Sang, M.K., Park, J.E., Song, D.K., Park, S.Y., Lee, J.D., Park, H.S., Jo, Y.H., Lee, J.S., Kang, S.W., and Lee, Y.S. (2020) Phylogenetic analysis of *Aegista quelpartensis* Using Metallothionein Gene. *The Korean Journal of Malacology,* **36**(2): 81-88.

Jeong, J.Y., Park, J.E., Song, D.K., Hong, C.E., Kim, Y.T., Shin, H.J., Liu, Z., Sang, M.K., Patnaik, H.H., Patnaik, B.B., Kang, S.W., Park, S.Y., Lee, J.S., Han, Y.S., Park, H.S., and Lee, Y.S. (2022) PANM DB ver 5.0 : An update of the PANM database for invertebrate NGS data analysis. *The Korean Journal of Malacology,* **38**(3): 125-128.

Jeong, J.Y., Sang, M.K., Park, J.E., Song, D.K., Hong, C.E., Kim, Y.T., Shin, H.J., Hwang, H.J., Jung, S.M., Park, S.Y., Kang, S.W., Lee, J.S., Han, Y.S., Park, H.S., Lee, Y.S., and Kim, W.J. (2021) Molecular Phylogenetic Study of *Clithon retropictus* using metallothionein gene. *The Korean Journal of Malacology,* **37**(2): 69-74.

Kang, S.W., Park, S., Y, Patnaik., B.B., Hwang, H.J., Kim, C.M., Kim, S.O., Lee, J.S., Han, Y.S., and Lee, Y.S. (2015) Construction of PANM Database (Protostome DB) for rapid annotation of NGS data in Mollusks. *The Korean Journal of Malacology,* **31**(3): 243-247.

Kang, S.W., Park, S.Y., Hwang, H.J., Chung, J.M., Sang, M.K., Min, H.R., Park, J.E., Cho, H.C., Patnaik, B.B., and Lee, Y.S. (2019) PANM DB ver 3.0 : An update of the bioinformatics database for annotation of large datasets from sequencing of species under Protostomia clade. *The Korean Journal of Malacology,* **35**(1): 73-75.

Kang, S.W., Patnaik, B.B., Hwang, H.J., Park, S.Y., Chung, J.M., Song, D.K., Patnaik, H.H., Lee, J.B., Kim, C., Kim, S., Park, H.S., Han, Y.S., Lee, J.S., and Lee, Y.S. (2016) Transcriptome sequencing and de novo characterization of Korean endemic land snail, *Koreanohadra kurodana* for functional transcripts and SSR markers. *Molecular Genetics and Genomics*., **291**(5): 1999-2014.

Kang, S.W., Patnaik, B.B., Hwang, H.J., Park, S.Y., Chung, J.M., Song, D.K., Patnaik, H.H., Lee, J.B., Kim, C., Kim, S., Park, H.S., Park, S.H., Park, Y.S., Han, Y.S., Lee, J.S., and Lee, Y.S. (2017) Sequencing and de novo assembly of visceral mass transcriptome of the critically endangered land snail *Satsuma myomphala*: Annotation and SSR discovery. *Comp Biochem Physiol Part D Genomics Proteomics,* **21**: 77-89.

Kang, S.W., Patnaik, B.B., Hwang, H.J., Park, S.Y., Wang, T.H., Park, E.B., Chung, J.M., Song, D.K., Patnaik, H.H., Lee, J.B., Kim, C., Kim, S., Park, H.S., Lee, J.S., Han, Y.S., and Lee, Y.S. (2016) De novo Transcriptome Generation and Annotation for Two Korean Endemic Land Snails, *Aegista chejuensis* and *Aegista quelpartensis*, Using Illumina Paired-End Sequencing Technology. *International Journal of Molecular Sciences*., **17**(3): 379.

Kang, S.W., Patnaik, B.B., Park, S.Y., Hwang, H.J., Chung, J.M., Sang, M.K., Min, H.R., Park, J.E., Seong, J., Jo, Y.H., Noh, M.Y., Lee, J.D., Jung, K.Y., Park, H.S., Han, Y.S., Lee, J.S., and Lee, Y.S. (2018) Transcriptome analysis of the threatened snail *Ellobium chinense* reveals candidate genes for adaptation and identifies SSRs for conservation genetics. *Genes Genomics,* **40**(4): 333-347.

Kim, K.H., and Harvell, C.D. (2004) The rise and fall of a six-year coral-fungal epizootic. *The American Naturalist*., **164**(5): S52-63.

Kohlmeyer, J., and Volkmann-Kohlmeyer, B. (1991) Illustrated Key to the Filamentous Higher Marine Fungi. *Botanica Marina.,* **34**: 1-61.

Martin, M. (2011) CUTADAPT removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal,* **17**(1)

Metzker, M.L. (2010) Sequencing technologies-the next generation. *Nature Reviews Genetics.* **11**: 31-46.

Min, H.R., Hwang, H.J., Chung, J.M., Sang, M.K., Cho, H.C., Park, J.E., Jung, K.Y., Park, H.S., Han, Y.S.,

and Lee, Y.S. (2018) Molecular Phylogenetic studies of *Mirus junensis* using Arginine kinase gene sequence. *The Korean Journal of Malacology,* **34**(2): 107-114.

Morozova, O., and Marra, M.A. (2008) Applications of next-generation sequencing technologies in functional genomics. *Genomics,* **92**: 255-264.

Nagahama, T. (2006) Yeast Biodiversity in Freshwater, Marine and Deep-Sea Environments, *Biodiversity and Ecophysiology of Yeasts*., pp. 241-262

Osman, K.T. (2013) Biological Properties of Forest Soils. ***In***; Osman KT (ed) *Forest Soils: Properties and Management. Springer International Publishing, Cham*., pp. 77-95

Park, J.E., Cho, H.C., Hwang, H.J., Chung, J.M., Sang, M.K., Min, H.R., Kang, S.W., Park, S.Y., Patnaik, B.B., Kim, W.J., Han, Y.S., Lee, J.S., and Lee, Y.S. (2018) Molecular Phylogenetics of Korean endemic land snail, *Aegista chejuensis* inferred from Metallothionein gene sequence. *The Korean Journal of Malacology,* **34**(1): 59-65.

Park, J.E., Sang, M.K., Hwang, H.J., Song, D.K., Jeong, J.Y., Park, S.Y., Kang, S.W., Park, H.S., Han, Y.S., Lee, J.S., and Lee, Y.S. (2020) Molecular Phylogenetic Analysis of Ellobium chinense Using Arginine Kinase. *The Korean Journal of Malacology,* **36**(3): 133-141.

Park, S.Y., Patnaik, B.B., Kang, S.W., Hwang, H.J., Chung, J.M., Song, D.K., Sang, M.K., Patnaik, H.H., Lee, J.B., Noh, M.Y., Kim, C., Kim, S., Park, H.S., Lee, J.S., Han, Y.S., and Lee, Y.S. (2016) Transcriptomic Analysis of the Endangered Neritid Species *Clithon retropictus*: De Novo Assembly, Functional Annotation, and Marker Discovery. *Genes* (Basel), **7**(7): 35

Patnaik, B.B., Chung, J.M., Hwang, H.-J., Sang, M., Park, J.E., Min, H., Cho, H., Dewangan, N., Baliarsingh, S., Kang, S.W., Park, S., Jo, Y.H., Park, H.-S., Kim, W., Han, Y., Lee, J., and Lee, Y.S. (2019) Transcriptome analysis of air-breathing land slug, Incilaria fruhstorferi reveals functional insights into growth, immunity, and reproduction. *BMC Genomics,* **20**

Patnaik, B.B., Wang, T.H., Kang, S.W., Hwang, H.J., Park, S.Y., Park, E.B., Chung, J.M., Song, D.K., Kim, C., Kim, S., Lee, J.S., Han, Y.S., Park, H.S., and Lee, Y.S. (2016) Sequencing, De Novo Assembly, and Annotation of the Transcriptome of the Endangered Freshwater Pearl Bivalve, *Cristaria plicata*, Provides Novel Insights into Functional Genes and Marker Discovery. *PLoS One,* **11**: e0148622.

Pertea, G., Huang, X., Liang, F., Antonescu, V., Sultana, R., Karamycheva, S., Lee, Y., White, J., Cheung, F., Parvizi, B., Tsai, J., and Quackenbush, J. (2003) TIGR Gene Indices clustering tools (TGICL): a software system for fast clustering of large EST datasets. *Bioinformatics,* **19**(5): 651-652.

Ramasamy, P., Vino, A.B., Saravanan, R., Subhapradha, N., Shanmugam, V., and Shanmugam, A. (2011) Screening of antimicrobial potential of polysaccharide from cuttlebone and methanolic extract from body tissue of Sepia prashadi Winkworth, 1936. *Asian Pacific Journal of Tropical Biomedicine,* **1**(2): S244-S248.

Ryu, Y.H. (2017) Cytochalasan and Azaphilone Derivatives from a Marin-Derived Fungus Chaetomium globosum

Sang, M.K., Hwang, H.J., Chung, J.M., Park, J.E., Song, D.K., Jeong, J.Y., Kang, S.W., Han, Y.S., Lee, Y.S., and Park, S.Y. (2020) Trends in genome and transcriptome research on mollusks in the world (2020). *The Korean Journal of Malacology,* **36**(1): 1-6.

Sang, M.K., Hwang, H.J., Chung, J.M., Park, J.E., Song, D.K., Jeong, J.Y., Park, S.Y., Park, H.S., Cho, Y.H., Lee, J.D., Lee, J.S., Lee, Y.S., and Kang, S.W. (2020) Phylogenetic analysis of endangered wild animal Class II Satsuma myomphala using the argrinine kinase gene. *The Korean Journal of Malacology,* **36**(2): 97-104.

Sang, M.K., Hwang, H.J., Kang, S.W., Park, S.H., Park, S.Y., Chung, J.M., Park, J.E., Min, H.R., Lee, J.S., Han, Y.S., Park, H.S., Won, R., and Lee, Y.S. (2017) Molecular phylogenetic study of *Incilaria fruhstorferi* based on Metallothionein gene. *The Korean Journal of Malacology,* **33**(4): 259-265.

Sang, M.K., Kang, S.W., Hwang, H.J., Chung, J.M., Song, D.K., Min, H.R., Park, J.E., Ha, H.C., Lee, H.J., Hong, C.E., Ahn, Y.M., Park, S.Y., Park, Y.S., Park, H.S., Han, Y.S., Lee, J.S., and Lee, Y.S. (2016) Molecular Phylogenetic Study of the Endangered Land Snail *Satsuma myomphala* Based on Metallothionein Gene. *The Korean Journal of Malacology,* **32**(4): 263-268.

Udo, I., Udoh, D.I., and Isang, O. (2022) Microbiological and Proximate Evaluation of Knife Clam (Tagelus Adansonii, Bosc, 1801) from Okoro River, Eastern Obolo, Nigeria. *Journal of Fisheries Science,* **4**(1)

van Dijk, E.L., Auger, H., Jaszczyszyn, Y., and Thermes, C. (2014) Ten years of next-generation sequencing technology. *Trends in Genetics,* **30**(9): 418-426.

Wickham, H. (2007) Reshaping Data with the reshape Package. *Journal of Statistical Software.* **21**(12): 1-20.

Wickham, H., François, R., Henry, L., and Müller, K. (2020) A Grammar of Data Manipulation [R package dplyr version 1.0.2]

Yang, R., Guo, X., Yang, J., Jiang, Y., Pang, B., Chen, C., Yao, Y., Qin, J., and Li, Q. (2009) Genomic research for important pathogenic bacteria in China. *Science in China Series C: Life Sciences,* **52**(1): 50-63.

Zhang, Y., Mu, J., Feng, Y., Kang, Y., Zhang, J., Gu, P.-J., Wang, Y., Ma, L.-F., and Zhu, Y.-H. (2009) Broad-Spectrum Antimicrobial Epiphytic and Endophytic Fungi from Marine Organisms: Isolation, Bioassay and Taxonomy. *Marine Drugs,* **7**(2): 97-112.

Zhu, C., Zhang, L., Ding, H., and Pan, Z. (2019) Transcriptome-wide identification and characterization of the Sox gene family and microsatellites for

Corbicula fluminea. *PeerJ.,* **7**: e7770.

Zvereva, L.V., and Vysotskaya, M.A. (2005) Filamentous Fungi Associated with Bivalve Mollusks from Polluted Biotopes of Ussuriiskii Bay, Sea of Japan. *Russian Journal of Marine Biology,* **31**: 382-385.